

1645

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/446,677

DATE: 09/06/2000  
 TIME: 11:01:37

Input Set : A:\Birkell.txt  
 Output Set: N:\CRF3\09062000\I446677.raw

Does Not Comply  
 Corrected Diskette Needed

2, 4<sup>6</sup>  
 See pp.

SEQUENCE LISTING  
 3 (1) GENERAL INFORMATION:  
 5 (i) APPLICANT: BIRKELUND, Svend  
 6 CHRISTIANSEN, Gunna  
 7 HEBSGAARD PEDERSEN, Anna-Sofie  
 8 MYGIND, Per  
 9 KNUDSEN, Katrine  
 11 (ii) TITLE OF INVENTION: SURFACE EXPOSED PROTEINS FROM CHLAMYDIA  
 12 PNEUMONIAE  
 14 (iii) NUMBER OF SEQUENCES: 30  
 16 (iv) CORRESPONDENCE ADDRESS:  
 17 (A) ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
 18 (B) STREET: 624 Ninth Street, N.W., Suite 300  
 19 (C) CITY: Washington  
 20 (D) STATE: D.C.  
 21 (E) COUNTRY: USA  
 22 (F) ZIP: 20001  
 24 (v) COMPUTER READABLE FORM:  
 25 (A) MEDIUM TYPE: Floppy disk  
 26 (B) COMPUTER: IBM PC compatible  
 27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
 30 (vi) CURRENT APPLICATION DATA:  
 C--> 31 (A) APPLICATION NUMBER: US/09/446,677  
 C--> 32 (B) FILING DATE: 24-Mar-2000  
 C--> 38 (vii) PRIOR APPLICATION DATA:  
 35 (A) APPLICATION NUMBER: PCT/DK98/00266  
 36 (B) FILING DATE: 19-JUN-1998  
 39 (A) APPLICATION NUMBER: DK 0744/97  
 40 (B) FILING DATE: 23-JUN-1997  
 42 (viii) ATTORNEY/AGENT INFORMATION:  
 43 (A) NAME: COOPER, Iver P.  
 44 (B) REGISTRATION NUMBER: 28,005  
 45 (C) REFERENCE/DOCKET NUMBER: BIRKELUND=1  
 47 (ix) TELECOMMUNICATION INFORMATION:  
 48 (A) TELEPHONE: 202-628-5197  
 49 (B) TELEFAX: 202-737-3528

## ERRORED SEQUENCES

498 (2) INFORMATION FOR SEQ ID NO: 3:  
 500 (i) SEQUENCE CHARACTERISTICS:  
 501 (A) LENGTH: 2815 base pairs  
 502 (B) TYPE: nucleic acid  
 503 (C) STRANDEDNESS: single  
 504 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/446,677

DATE: 09/06/2000  
TIME: 11:01:37

Input Set : A:\Birkell.txt  
Output Set: N:\CRF3\09062000\I446677.raw

Enter "hard return"  
here to correct.

506 (ii) MOLECULE TYPE: Genomic DNA  
508 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
E--> 510 ATGAAATCGC AATTTTCTG GTTAGGCTC TCTTCGACAT TGGCATGTTT TACTACTTGT 60 TCCACTGTTT TTGCTGCAAC TGCTGAAAT ATA  
511 ACTAACACAG GCACCTATAC TCCTAAAAAT ACGAACATG GAATAGACTA TACTCTGACA 180  
512 GGAGATATAA CTCTGCAAA CCTCTGGGAT TCAGCAGCTT TAACGAAGGG TTGTTTTCT 240  
513 GACACTACGG AATCTTAAAG CTTTCCGGT AAGGGTACT CACTTCCTT TTAAATATT 300  
514 AAGTCTAGT CTGAAGGCAG AGCCTTCTT GTTACAACATG ATAAAATCT 360  
515 GGATTTCGA GTCTTACTTT CTTAGCGGCC CCATCATCGG TAATCACAAC CCCCTCAGGA 420  
516 AAAGGTGAG TAAATGTGG AGGGATCTT ACATTTGATA ACAATGGAAC TATTTTATT 480  
517 AAACAAGATT ACTGTGAGGA AAATGGCGGA GCCATTCTA CCAAGAATCT TTCTTGAAA 540  
518 AACAGCACGG GATCGATTTC TTTTGAAAGG AATAATCGA GCGCAACAGG GAAAAAAGGT 600  
519 GGGGCTATT GTGCTACTGG TACTGTAGAT ATTACAAATA ATACGGCTCC TACCCCTTTC 660  
520 TCGAACATA TTGCTGAAGC TGCAGGTGGA GCTATAATA GCACAGGAA CTGTACAATT 720  
521 ACAGGGAATA CGTCTCTTGT ATTTTCTGAA AATAGTGTGA CAGCAGCCGC AGGAAATGGA 780  
522 CGAGCTCTT CTGGAGATGC CGATGTTTAC ATATCTGGGA ATCAGAGTGT AACTTTCTCA 840  
523 CGAACCAAG CTGTAGCTAA TGGCGGAGCC ATTATGCTA AGAAGCTTAC ACTGGCTTCC 900  
524 GGGGGGGGGG GGGGTATCTC CTTTCTAAC AATATAGTCC AAGGTACCC TGCAGGTAAT 960  
525 GGTGGAGCCA TTTCTATACT GGCAGCTGGA GACTGTAGTC TTTCAGCAGA ACCAGGGGAC 1020  
526 ATTACCTCA ATGGGAATGC CATTGTTGCA ACTACACCAC AAACATACAA AGAAATTCT 1080  
527 ATTGACATAG GATCTACTGC AAAGATCACC AATTTCAGTG CAATATCTGG GCATAGCATC 1140  
528 TTTTCTACG ATCCGATTAC TCCTAAATCG GCTGCGGATT CTACAGATAC TTAAATCTC 1200  
529 AATAAGGCTG ATGCAGGTA TAGTACAGAT TATAGTGGGT CGATTGTTT TTCTGGTGA 1260  
530 AAGCTCTCG AAGATGAAGC AAAAGTTGCA GACAACCTCA CTTCTACGCT GAAGCAGCCT 1320  
531 GTAACTCTAA CTGCAAGGAAA TTATGACTT AAACGTTGTC TCACTCTCGA TACGAAAGC 1380  
532 TTACTCAGA CGCGGGGTTCTCTGTATT ATGGATGCGG GCACAACTGTT AAAAGCAAGT 1440  
533 ACAGAGGAGG TCACTTTAAC AGGTCTTCC ATTCCTGTAG ACTCTTTAGG CGAGGGTAAG 1500  
534 AAAGTTGAA TTGCTGCTTC TGCAAGCAAGT AAAATGTA CGCTTAGTGG TCCGATTCTT 1560  
535 CTTTTGGATA ACCAAGGGA TGCTTATGAA AATCAGACT TAGGAAAAAC TCAAGACTTT 1620  
536 TCATTTGTGC AGCTCTCTG TCTGGTACT GCAACAACTA CAGATGTTCC AGCGGTTCT 1680  
537 ACAGTACCAA CTCTCTAGCA CTATGGGTAT CAAGGTTACTT GGGGAATGAC TTGGGTTGAT 1740  
538 GATACCGCAA GCACCTCAA GACTAAGACA GGCACATAG CTTGGACCAA TACAGCTAC 1800  
539 CTTCCGAATC CTGAGCGTCA AGGACCTTTA GTTCTTAATA GCTTGTGGG ATCTTTTCA 1860  
540 GACATCCAAG CGATTCAAGG TGTCTAGAG AGAAGTGTGTT TGACTCTTTC TTCAAGATCGA 1920  
541 GGCTTCTGGG CTGCGGGAGT CGCCAATTTC TTAGATAAAG ATAAGAAAGG GGAAAAAACGC 1980  
542 AAATACCGTC ATAATCTGG TGGATATGCT ATCGGAGGTC CAGCGCAACAC TTGTTCTGAA 2040  
543 AACTTAATTG GCTTGTGCTT TTGCTCAACTC TTTGGTAGCG ATAAAGATT CTTAGTCCT 2100  
544 AAAATCTATA CTGATACCTA TGCAAGGAGCC TTCTATATCC AACACATTAC AGAATGTAGT 2160  
545 GGGTTCTACATG GTTGTCTCTT AGATAAACCTT CCTGGCTCTT GGAGTCATAA ACCCCTCTT 2220  
546 TTAGAAGGGC AGCTGCTTA TAGCCACGTC AGTAATGATC TGAAGACAAA GTATACTGG 2280  
547 TATCCTGAGG TGAAAGGTTT TTGGGGAAAT AATGCTTTTA ACATGATGTT GGGAGCTCT 2340  
548 TCTCATCTT ATCCGAAATA CCTGCAACTC TTTGGTAGCG ATGCTCCATA CATCAAACCTG 2400  
549 AATCTGACCT ATATACGTC GCAAGCTTC TCGGAGAAAG GTACAGAAGG AAGATCTTT 2460  
550 GATCACAGCA ACCTCTTCAA TTATCTTTC CCTATAGGGG TGAAGTTGGA GAAGTTCTCT 2520  
551 GATTTGAAATG ACTTTCTTA TGATCTGACT TTATCTTATG TTCTGTGATCT TATCCGAAT 2580  
552 GATCCCAAAT GCACATACAGC ACTTGAAATC AGCGGAGCCT CTTGGGAAAC TTATGCAAT 2640  
553 AACTTAGCAC GACAGGGCTT GCAAGTGCCT GCAAGGAGTC ACTACGCTT CTCTCCTATG 2700  
554 TTTGAAGTGC TCGGCCACTT TGTCTTGAA GTTCGTGGAT CCTCACGGAT TTATAATGTA 2760  
555 GATCTGGGG GTAAGTTCCA ATTCTAGGAG CGTCTCTCAT GTCTCAGAAA TTCTG 2815

743 (2) INFORMATION FOR SEQ ID NO: 5:

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/446,677

DATE: 09/06/2000  
TIME: 11:01:37

Input Set : A:\Birkell.txt  
Output Set: N:\CRF3\09062000\I446677.raw

745 (i) SEQUENCE CHARACTERISTICS:  
746 (A) LENGTH: 3052 base pairs → See p. 4  
747 (B) TYPE: nucleic acid  
748 (C) STRANDEDNESS: single  
749 (D) TOPOLOGY: linear  
751 (ii) MOLECULE TYPE: Genomic DNA  
753 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
755 ATGCGATTT CGCTCGCG ATTCTCTTA GTTTTTCTT TAACATTGCT CTCAGTCTTC 60  
756 GACACTTCT TGAGTGTAC TACGATTCT TTAACCCAG AGATAGTTT TCATGGAGAT 120  
757 AGTCAGAATG CAGAACGTT TTATAATGTT CAAGCTGGGG ATGCTATAG CCTTACTGGT 180  
758 GATGTCCTAA TATCTAAGCT CGATAACTCT GCATTAATAA AGCCCTGCTT CAATGTGACC 240  
759 TCAAGGAAGTG TGACGTCGC AGGAAATCAT CATGGTTAT ATTCTAATAA TATTCCTCA 300  
760 GGAACATCAA AGGAAGGGGC TGACTTTGT TGCCAAGATC CTCAAGCAAC GGCACGTTT 360  
761 TCTGGTTCT CCACGCTCTC TTTTATTAG AGCCCCGGAG ATATTTAAAGA ACAGGGATGT 420  
762 CTCTATTCAA AAAATGACTT TATGCTCTTA AACATTATG TAGTGCCTT TGAACAAAC 480  
763 CAAAGTAAGA CTAAAGGGG AGCTTATTAGT GGGCGGAATG TTACTATAGT AGGCAACTAC 540  
764 GATTCGGTCT CTTCTATCA GAATGCAGCC ACTTTTGGAG GTGCTATCCA TTCTTCAGGT 600  
765 CCCCTACAGA TTGCACTAA TCAGGCAGAG ATAAGATTG CACAAATAC TGCCAAGAAAT 660  
766 GGTTCTGGAG GGGCTTGTA CTCCGATGGT GATATTGATA TTGATCAGAA TGCTTATGTT 720  
767 CTATTCGAG AAAATGAGGC ATTGACTACT GCTATGGTA AGGGAGGGC TGCTGTTT 780  
768 CTTCCCACTT CAGGAAGTAG TACTCCAGTT CCTATTGTA CTTTCTCTGA CAATAAACAG 840  
769 TTAGCTTTG AAAGAAACCA TTCCATAATG GTGCGCGAG CCATTTATGC TAGGAAACTT 900  
770 AGCATCTCTT CAGGAGGTCC TACTCTATTG ATCAATAATA TATCATATGC AAATTCGCAA 960  
771 AATTTAGGTG GAGCTATTGC CATTGATACT GGAGGGAGA TCAGTTTATC AGCAGAGAAA 1020  
772 GGAACAAATTA CATTCCAAGG AAACGGGACG AGCTTACCGT TTTTGAATGG CATCCATCTT 1080  
773 TTACAAAATG CTAAATTCTT GAAATTACAG GCGAGAAATG GATGCTCTAT AGAATTTTAT 1140  
774 GATCCTATTA CTTCTGAAGC AGATGGGTCT ACCAATTGTA ATATCAACGG AGATCCTAAA 1200  
775 AATAAAAGAGT ACACAGGGAC CATACTCTT TCTGGAGAAA AGAGTCTAGC AAACGATCCT 1260  
776 AGGGATTATA AATCTACAACT CCCCTACAGA AC GTCACCTGT CTGAGGATA CTTAGTTATT 1320  
777 AAAGAGGGGG CGGAAGTCAC AGTTTCAAAA TTACACGAGT CTCCAGGATC GCATTTAGTT 1380  
778 TTAGATTAGG CAACCAAATC GATAGCCTCT AAGGAAGACA TTGCCATCAC AGGCCCTCGG 1440  
779 ATAGATATAG ATAGCTTAAG CTCATCTCA ACAGCAGCTG TTATTAAGC AAACACCGCA 1500  
780 AATAAAACAGA TATCCGTGAC GGACTCTATA GAACCTATGC CGCTTACTGG CAATGCCAT 1560  
781 GAAGATCTCA GAATGAGAAA TTACACAGC TTCCCTCTGC TCTCTTTAGA GCCTGGAGCC 1620  
782 GGGGGTAGTG TGACTGTAAC TGCTGGAGAT TTCTTACCGG TAAGTCCCCA TTATGTTT 1680  
783 CAAGGCAATT GGAATTAGC TTGGACAGGA ACTGGAAACA AAGTTGGAGA ATTCTTCTGG 1740  
784 GATAAAATAA ATTATAAGC TAGACCTGAA AAAAGAAGGA ATTTAGTTCC TAATATCTG 1800  
785 TCGGGGAATG CTGTAATGT CAGATCTTA ATGCAGGTTG AAGAGACCCA TGCATCGAGC 1860  
786 TTACAGACAG ATCGAGGGCT GTGGATCGAT GGAATTGGGA ATTTCTTCCA TGTATCTGCC 1920  
787 TCCGAAGACA ATATAAGGT CCGTCATAAC AGCGGTGGAT ATGTTCTATC TGTAATAAT 1980  
788 GAGATCACAC CTAACCACTA TACTTCGATG GCATTTTCCC AACTCTTTAG TAGAGACAG 2040  
789 GACTATGCGG TTTCACAAAGA CGAATACAGA ATGTATTTAG GATGCTATCT CTATCAATAT 2100  
790 ACAACCTCCC TAGGGAATAT TTTCGTTAT GCTTCGCGTA ACCCTAATGT AAACGTCGGG 2160  
791 ATTCTCTCAA GAAGGTTCTC TCAAAATCTT CTTATGATT TTCACTTTTT GTGTGCTTAT 2220  
792 GGTGATGCCA CCAATGATAT GAAAACAGAC TACGCCAAATT TCCCTATGGT GAAAAAACAGC 2280  
793 TGGAGAAACA ATTGTTGGC TATAGAGTC GGAGGGAGCA TGCCCTTATT GGTATTGAG 2340  
794 AACGGAAGAC TTTTCCAAGG TGCCATCCCA TTATGAAAC TACAATTAGT TTATGCTTAT 2400  
795 CAGGGAGATT TCAAAGAGAC GACTGCAGAT GGCCGTAGAT TTAGTAATGG GAGTTAACAA 2460  
796 TCGATTCTCG TACCTCTAGG CATAACGCTTT GAGAAGCTGG CACTTTCTCA GGATGTACTC 2520

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/446,677

DATE: 09/06/2000  
TIME: 11:01:37

Input Set : A:\Birkell.txt  
Output Set: N:\CRF3\09062000\I446677.raw

797	TATGACTTTA	GTTTCTCCTA	TATTCTCTGAT	ATTTCCGTA	AGGATCCCTC	ATGTGAAGCT	2580
798	GCTCTGGTGA	TTAGCGGAGA	CTCCTGGCTT	GTTCCGGCAG	CACACCTATC	AAGACATGCT	2640
799	TTTGTAGGATAG	GTGGAACGGG	TCGGTATCAC	TTTAACGACT	ATACTGAGCT	CTTATGTCGA	2700
800	GGAAAGTATAG	AATGCCGCC	CCATGCTAGG	AATTATAATA	TAAGACTGTGG	AAGCAAATTT	2760
801	CGTTTTAGA	AGGTTTCCAT	TGCCTGTGTG	GTTCCGGATC	TTAACATATA	ATCCTGGACT	2820
802	ATGGATCATCA	GGCATTGGGT	TTCTCGAATC	TGTGTGGAGA	ATAACGACAT	TTTATATGCA	2880
803	TAACCGAATA	CTCGTATCAC	CTCAGCCCT	AGAGACATT	TTTAGGGTT	CTTATTTGT	2940
804	CTAAACCTTCG	TATTTATCG	AGAACCTTTT	ACGTTCTTGG	TTTGCTTGTC	TCCGAGGAGT	3000
E--> 805	TCTCTAACGA	ATCATAGGGA	TTCCAGGGTT	CTGTTCCCTG	AGTCCCTTGG	A	3052
2597	(2) INFORMATION FOR SEQ ID NO: 24:						
2599	(i) SEQUENCE CHARACTERISTICS:						
2600	(A) LENGTH: <u>946</u> amino acids						→ See p. 6
2601	(B) TYPE: amino acid						
2602	(C) STRANDEDNESS: single						
2603	(D) TOPOLOGY: linear						
2605	(ii) MOLECULE TYPE: peptide						
2607	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:						
2609	Met Lys Thr Ser Val Ser Met Leu Leu Ala Leu Leu Cys Ser Gly Ala						
2610	1	5	10	15			
2612	Ser Ser Ile Val Leu His Ala Ala Thr Thr Pro Leu Asn Pro Glu Asp						
2613	20	25	30				
2615	Gly Phe Ile Gly Glu Gly Asn Thr Asn Thr Phe Ser Pro Lys Ser Thr						
2616	35	40	45				
2618	Thr Asp Ala Ala Gly Thr Thr Tyr Ser Leu Thr Gly Glu Val Leu Phe						
2619	50	55	60				
2621	Ile Asp Pro Gly Lys Gly Ser Ile Thr Gly Thr Cys Phe Val Glu						
2622	65	70	75	80			
2624	Thr Ala Gly Asp Leu Thr Phe Leu Gly Asn Gly Asn Thr Leu Lys Phe						
2625	85	90	95				
2627	Leu Ser Val Asp Ala Gly Ala Asn Ile Ala Val Ala His Val Gln Gly						
2628	100	105	110				
2630	Ser Lys Asn Leu Ser Phe Thr Asp Phe Leu Ser Leu Val Ile Thr Glu						
2631	115	120	125				
2633	Ser Pro Lys Ser Ala Val Ser Thr Gly Lys Gly Ser Leu Val Ser Ser						
2634	130	135	140				
2636	Gly Ala Val Gln Leu Gln Asp Ile Asn Thr Leu Val Leu Thr Ser Asn						
2637	145	150	155	160			
2639	Ala Ser Val Glu Asp Gly Gly Val Ile Lys Gly Asn Ser Cys Leu Ile						
2640	165	170	175				
2642	Gln Gly Ile Lys Asn Ser Ala Ile Phe Gly Gln Asn Thr Ser Ser Lys						
2643	180	185	190				
2645	Lys Gly Gly Ala Ile Ser Thr Thr Gln Gly Leu Thr Ile Glu Asn Asn						
2646	195	200	205				
2648	Leu Gly Thr Leu Lys Phe Asn Glu Asn Lys Ala Val Thr Ser Gly Gly						
2649	210	215	220				
2651	Ala Leu Asp Leu Gly Ala Ala Ser Thr Phe Thr Ala Asn His Glu Leu						
2652	225	230	235	240			
2654	Ile Phe Ser Gln Asn Lys Thr Ser Gly Asn Ala Ala Asn Gly Gly Ala						
2655	245	250	255				

Number of bases  
conflict, 3052  
listed, 3051. found.

RECEIVED

TECH CENTER 1600, 2000

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/446,677DATE: 09/06/2000  
TIME: 11:01:38Input Set : A:\Birkell.txt  
Output Set: N:\CRF3\09062000\I446677.raw

2657 Ile Asn Cys Ser Gly Asp Leu Thr Phe Thr Asp Asn Thr Ser Leu Leu  
 2658 260 265 270  
 2660 Leu Gln Glu Asn Ser Thr Met Gln Asp Gly Gly Ala Leu Cys Ser Thr  
 2661 275 280 285  
 2663 Gly Thr Ile Ser Ile Thr Gly Ser Asp Ser Ile Asn Val Ile Gly Asn  
 2664 290 295 300  
 2666 Thr Ser Gly Gln Lys Gly Gly Ala Ile Ser Ala Ala Ser Leu Lys Ile  
 2667 305 310 315 320  
 2669 Leu Gly Gly Gln Gly Gly Ala Leu Phe Ser Asn Asn Val Val Thr His  
 2670 325 330 335  
 2672 Ala Thr Pro Leu Gly Gly Ala Ile Phe Ile Asn Thr Gly Gly Ser Leu  
 2673 340 345 350  
 2675 Gln Leu Phe Thr Gln Gly Gly Asp Ile Val Phe Glu Gly Asn Gln Val  
 2676 355 360 365  
 2678 Thr Thr Ala Pro Asn Ala Thr Thr Lys Arg Asn Val Ile His Leu  
 2679 370 375 380  
 2681 Glu Ser Thr Ala Lys Trp Thr Gly Leu Ala Ala Ser Gln Gly Asn Ala  
 2682 385 390 395 400  
 2684 Ile Tyr Phe Tyr Asp Pro Ile Thr Thr Asn Asp Thr Gly Ala Ser Asp  
 2685 405 410 415  
 2687 Asn Leu Arg Ile Asn Glu Val Ser Ala Asn Gln Lys Leu Ser Gly Ser  
 2688 420 425 430  
 2690 Ile Val Phe Ser Gly Glu Arg Leu Ser Thr Ala Glu Ala Ile Ala Glu  
 2691 435 440 445  
 2693 Asn Leu Thr Ser Arg Ile Asn Gln Pro Val Thr Leu Val Glu Gly Ser  
 2694 450 455 460  
 2696 Leu Glu Leu Lys Gln Gly Val Thr Leu Ile Thr Gln Gly Phe Ser Gln  
 2697 465 470 475 480  
 2699 Glu Pro Glu Ser Thr Leu Leu Leu Asp Leu Gly Thr Ser Leu Gln Ala  
 2700 485 490 495  
 2702 Ser Thr Glu Asp Ile Val Ile Thr Asn Ser Ser Ile Asn Ala Asp Thr  
 2703 500 505 510  
 2705 Ile Tyr Gly Lys Asn Pro Ile Asn Ile Val Ala Ser Ala Ala Asn Lys  
 2706 515 520 525  
 2708 Asn Ile Thr Leu Thr Gly Thr Leu Ala Leu Val Asn Ala Asp Gly Ala  
 2709 530 535 540  
 2711 Leu Tyr Glu Asn His Thr Leu Gln Asp Ser Gln Asp Tyr Ser Phe Val  
 2712 545 550 555 560  
 2714 Lys Leu Ser Pro Gly Ala Gly Gly Thr Ile Ile Thr Gln Asp Ala Ser  
 2715 565 570 575  
 2717 Gln Lys Leu Leu Glu Val Ala Pro Ser Arg Pro His Tyr Gly Tyr Gln  
 2718 580 585 590  
 2720 Gly His Trp Asn Val Gln Val Ile Pro Gly Thr Gly Thr Gln Pro Ser  
 2721 595 600 605  
 2723 Gln Ala Asn Leu Glu Trp Val Arg Thr Gly Tyr Leu Pro Asn Pro Glu  
 2724 610 615 620  
 2726 Arg Gln Gly Phe Leu Val Pro Asn Ser Leu Trp Gly Ser Phe Val Asp  
 2727 625 630 635 640  
 2729 Gln Arg Ala Ile Gln Glu Ile Met Val Asn Ser Ser Gln Ile Leu Cys

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/446,677

DATE: 09/06/2000  
TIME: 11:01:38

Input Set : A:\Birkell.txt  
Output Set: N:\CRF3\09062000\I446677.raw

2730	645	650	655
2732	Gln Glu Arg Gly Val Trp Gly Ala Gly	Ile Ala Asn Phe Leu His Arg	
	660	665	670
2733	Asp Lys Ile Asn Glu His Gly Tyr Arg His Ser Gly Val	Gly Tyr Leu	
	675	680	685
2736	Val Gly Val Gly Thr His Ala Phe Ser Asp Ala Thr	Ile Asn Ala Ala	
	690	695	700
2738	705	710	715
2739	Phe Cys Gln Leu Phe Ser Arg Asp Lys Asp Tyr Val Val Ser Lys Asn		720
2741	725	730	735
2742	His Gly Thr Ser Tyr Ser Gly Val Val Phe Leu Glu Asp Thr Leu Glu		
2744	740	745	750
2745	Phe Arg Ser Pro Gln Gly Phe Tyr Thr Asp Ser Ser Ser Glu Ala Cys		
2747	755	760	765
2748	Cys Asn Gln Val Val Thr Ile Asp Met Gln Leu Ser Tyr Ser His Arg		
2750	770	775	780
2751	Asn Asn Asp Met Lys Thr Lys Tyr Thr Thr Tyr Pro Glu Ala Gln Gly		
2753	785	790	795
2754	Ser Trp Ala Asn Asp Val Phe Gly Leu Glu Phe Gly Ala Thr Thr Tyr		800
2756	805	810	815
2757	Tyr Tyr Pro Asn Ser Thr Phe Leu Phe Asp Tyr Tyr Ser Pro Phe Leu		
2759	820	825	830
2760	Arg Leu Gln Cys Thr Tyr Ala His Gln Glu Asp Phe Lys Glu Thr Gly		
2762	835	840	845
2763	Gly Glu Val Arg His Phe Thr Ser Gly Asp Leu Phe Asn Leu Ala Val		
2765	850	855	860
2766	Pro Ile Gly Val Lys Phe Glu Arg Phe Ser Asp Cys Lys Arg Gly Ser		
2768	865	870	875
2769	Tyr Glu Leu Thr Leu Ala Tyr Val Pro Asp Val Ile Arg Lys Asp Pro		880
2771	885	890	895
2772	Lys Ser Thr Ala Thr Leu Ala Ser Gly Ala Thr Trp Ser Thr His Gly		
2774	895	900	910
2775	Asn Asn Leu Ser Arg Gln Gly Leu Gln Leu Arg Leu Gly Asn His Cys		
2777	905	910	915
2778	Leu Ile Asn Pro Gly Ile Glu Val Phe Ser His Gly Ala Ile Glu Leu		
2780	920	925	930
2781	Arg Gly Ser Ser Arg Asn Tyr Asn Ile Asn Leu Gly Gly Lys Tyr Arg		
2783	935	940	945
2784	Phe		

E--> 2787 945

→ 945 amino acids found, 946 listed as length.

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/446,677

DATE: 09/06/2000  
TIME: 11:01:39

Input Set : A:\Birkell.txt  
Output Set: N:\CRF3\09062000\I446677.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:38 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]  
L:510 M:320 E: (1) Wrong Nucleic Acid Designator, 2  
L:805 M:254 E: No. of Bases conflict, Input:3052 Counted:3051 SEQ:5  
L:805 M:204 E: No. of Bases differ, LENGTH:Input:3052 Counted:3051 SEQ:5  
L:2454 M:220 C: Keyword misspelled or invalid format, [(ii) MOLECULE TYPE:]  
L:2787 M:203 E: No. of Seq. differs, LENGTH:Input:946 Found:945 SEQ:24